

1. A method of summarizing a text, comprising the steps of:
  - (a) receiving the text, where the text includes at least one set of textual units, where each set of textual units includes at least one textual unit;
  - (b) identifying all of the textual units in the text;
  - (c) selecting a first set of textual units from the text;
  - (d) selecting a second set of textual units from the text;
  - (e) identifying each unique textual unit in the first set of textual units;
  - (f) identifying each unique textual unit in the second set of textual units;
  - (g) determining how many textual units are shared between the first set of textual units and the second sets of textual units;
  - (h) selecting a third set of textual units from the text, where the third set of textual units is between the first set of textual units and the second set of textual units;
  - (i) identifying each unique textual unit in the third set of textual units;
  - (j) identifying each unique textual unit in the text;
  - (k) determining the frequency of occurrence of each unique textual unit in the third set of textual units;
  - (l) determining the frequency of occurrence of each unique textual unit in the text;
  - (m) determining the proximity of the results of step (k) and step (l);
  - (n) calculating a score for the first set of textual units with respect to the second set of textual units as a function of the results of step (g) and step (m);
  - (o) returning to step (d) if additional processing is desired, otherwise, proceeding to step (p);
  - (p) assigning the highest scoring result of step (n) to the first set of textual units;

(q) returning to step (c) if additional processing is desired, otherwise, proceeding to step (r); and

(r) selecting a user-definable number of first sets of textual units selected in step (c) according to the scores assigned thereto as the summary of the text.

2. The method of claim 1, wherein the step of selecting a second set of textual units from the text is comprised of the step of selecting a second set of textual units that occurs after the first set of textual units.

3. The method of claim 1, wherein the step of selecting a second set of textual units from the text is comprised of the step of selecting a second set of textual units, where the second set of textual includes at least one textual unit from the first set of textual units.

4. The method of claim 1, wherein the step of selecting a third set of textual units from the text is comprised of the step of selecting a third set of textual units that includes the last textual unit of the first set of textual units and the first textual unit of the second set of textual units.

5. The method of claim 1, wherein the step of determining proximity of the results of step (k) and step (l) is comprised of the steps of:

(a) multiplying, for each unique textual unit in the third set of textual units, the frequency of occurrence of the unique textual unit in the third set of textual units by a logarithm of the frequency of occurrence of the unique textual unit in the third set of textual units;

- (b) summing the results of the last step;
- (c) multiplying, for each unique textual unit in the third set of textual units, the frequency of occurrence of the unique textual unit in the third set of textual units by the logarithm of the frequency of occurrence of the unique textual unit in the text;
- (d) summing the results of the last step; and
- (e) dividing the result of step (b) by the result of step (d).

6. The method of claim 1, wherein the step of calculating a score for the first set of textual units with respect to the second set of textual units as a function of the results of step (g) and step (m) is comprised of the step of calculating the product of step (g) and step (m).

7. A method of summarizing a text, comprising the steps of:

- (a) receiving the text, where the text includes at least one set of textual units, where each set of at least one textual units includes at least one textual unit;
- (b) identifying each set of textual units in the text;
- (c) selecting a first set of textual units from the text;
- (d) selecting a second set of textual units from the text;
- (e) identifying each unique textual unit in the first set of textual units;
- (f) identifying each unique textual unit in the second set of textual units;
- (g) determining how many textual units are shared between the results of step (e) and step (f);
- (h) selecting a third set of textual units from the text, where the third set of textual units is between the first set of textual units and the second set of textual units;
- (i) identifying each unique textual unit in the third set of textual units;

- (j) identifying each unique textual unit in the text;
  - (k) determining the frequency of occurrence of each unique textual unit in the third set of textual units;
  - (l) determining the frequency of occurrence of each unique textual unit in the text;
  - (m) determining the proximity of the results of step (k) and step (l);
  - (n) calculating a score for the first set of textual units with respect to the second set of textual units as a function of the results of step (g) and step (m);
  - (o) returning to step (d) if additional processing is desired, otherwise, proceeding to step (p);
  - (p) assigning the highest result of step (n) to the first set of textual units;
  - (q) selecting a fourth set of textual units from the text, where the fourth set of textual units is contiguous with the first set of textual units;
  - (r) identifying each unique textual unit in the fourth set of textual units;
  - (s) determining how many textual units are shared between the results of step (e) and step (r);
  - (t) determining the frequency of occurrence of each unique textual unit in the fourth set of textual units;
  - (u) determining the proximity of the results of step (l) and step (t);
  - (v) calculating a score for the first set of textual units with respect to the fourth set of textual units as a function of the results of step (s) and step (u);
  - (w) returning to step (q) if additional processing is desired, otherwise, proceeding to step (x);
  - (x) combining a user-definable number of results of step (v) with the result of step (p);
  - (y) returning to step (c) if additional processing is desired, otherwise, proceeding to step (z);
- and

- (z) selecting a user-definable number of first set of textual units selected in step (c) according to the scores assigned thereto as the summary of the text.

8. The method of claim 7, wherein said step of calculating a score for the first set of textual units with respect to the fourth set of textual units as a function of the results of step (s) and step (u) is comprised of the steps of:

- (a) subtracting the result of step (u) from a number having a value equal to one; and
- (b) multiplying the result of the last step by the result of step (s).

9. The method of claim 7, wherein said step of combining a user-definable number of results of step (v) with the result of step (p) is comprised of the step of combining a user-definable number of results of step (v) with the results of step (p), where the results selected from step (v) are the highest values calculated in step (v).

10. A method of summarizing text, comprising the steps of:

- (a) receiving text, where the text includes at least one set of textual units, where each set of textual units includes at least one textual unit;
- (b) identifying the sets of textual units in the text;
- (c) selecting a first set of textual units from the text;
- (d) selecting a second set of textual units from the text;
- (e) identifying each unique textual unit in the first set of textual units;
- (f) identifying each unique textual unit in the second set of textual units;
- (g) determining how many textual units are shared between the results of step (e) and step (f);

- (h) selecting a third set of textual units from the text, where the third set of textual units is between the first set of textual units and the second sets of textual units;
- (i) identifying each unique textual unit in the third set of textual units;
- (j) identifying each unique textual unit in the text;
- (k) determining the frequency of occurrence of each unique textual unit in the third set of textual units;
- (l) determining the frequency of occurrence of each unique textual unit in the text;
- (m) determining the proximity of the results of step (k) and step (l);
- (n) selecting a fifth set of textual units from the text, where the fifth set of textual units is contiguous with the second set of textual units;
- (o) identifying each unique textual unit in the fifth set of textual units;
- (p) determining the frequency of occurrence of each unique textual unit in the fifth set of textual units;
- (q) determining the proximity of the results of step (l) and step (p);
- (r) combining the results of step (g), step (m), and step (q);
- (s) returning to step (d) if additional processing is desired, otherwise, proceeding to step (t);
- (t) assigning the highest result of step (n) to the first set of textual units;
- (u) returning to step (c) if additional processing is desired, otherwise, proceeding to step (v);
- and
- (v) selecting a user-definable number of first sets of textual units selected in step (c) according to the scores assigned thereto as the summary of the text.

11. A method of summarizing text, comprising the steps of:

- (a) receiving the text, where the text includes at least one set of textual units, where each set of textual units includes a plurality of contiguous textual units;
- (b) identifying the sets of textual units in the text;
- (c) identifying each unique textual unit in the text;
- (d) assigning a user-definable weight to each textual unit identified in step (c);
- (e) selecting a first set of textual units from the text;
- (f) selecting a second set of textual units from the text;
- (g) identifying each unique textual unit in the first set of textual units;
- (h) identifying each unique textual unit in the second set of textual units;
- (i) summing the weights of the textual units that are shared between the results of the step (g) and step (h);
- (j) selecting a third set of textual units from the text, where the third set of textual units is between the first set of textual units and the second set of textual units;
- (k) identifying each unique textual unit in the third set of textual units;
- (l) determining the frequency of occurrence of each unique textual unit in the third set of textual units;
- (m) determining the frequency of occurrence of each unique textual unit in the text;
- (n) determining the proximity of the results of step (l) and step (m);
- (o) calculating a score for the first set of textual units with respect to the second set of textual units as a function of the results of step (i) and step (n);
- (p) returning to step (f) if additional processing is desired, otherwise, proceeding to step (q);
- (q) assigning the highest result of step (o) to the first set of textual units;
- (r) returning to step (e) if additional processing is desired, otherwise, proceeding to step (s);

- 2044746-0140
- (s) selecting one of the unique textual units in the text;
  - (t) identifying, for the unique textual unit selected in step (s), each set of textual units in the text in which the selected unique textual unit appears for which a score was calculated, the score corresponding to each identified set of textual units, and the length of each identified set of textual units;
  - (u) summing the scores identified in step (t);
  - (v) recalculating the weight of the unique textual unit selected in step (s) as the combination of the result of step (u) and the lengths and the scores identified in step (t);
  - (w) returning to step (s) if additional unique textual units are desired to be weighted, otherwise, proceeding to step (x);
  - (x) if additional processing is desired, returning to the step (e), otherwise, proceeding to step (y);
  - (y) assigning the highest result of step (o) to the first set of textual units;
  - (z) returning to step (e) if additional processing is desired, otherwise, proceeding to step (aa);
  - and
  - (aa) selecting a user-definable number of first sets of textual units selected in step (e) according to the scores assigned thereto as the summary of the text.

12. The method of claim 11, further comprising the steps of:

- (a) setting a user-definable stop-word threshold;
- (b) identifying each textual unit having a weight below the stop-word threshold; and
- (c) removing the textual units identified in the last step from the text.



13. The method of claim 11, further comprising the steps of:

- (a) setting a user-definable key-word threshold;
- (b) identifying each textual units having a weight above the key-word threshold; and
- (c) returning a key-word summary of the text that includes the textual units identified in the last step.